

SEQUENCE LISTING

<110> BENOIT, Patrick
SCHWARTZ, Bertrand
BRANELLEC, Didier
CHIEN, Kenneth R.

<120> SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
THEM AND USES THEREOF

<130> 03806.0530-00000

<140>

<141> 2001-12-07

<160> 5

<170> PatentIn Ver. 2.1

<210> 1

<211> 2358

<212> DNA

<213> Mus musculus

<400> 1

gggatcctttc atgttaaca atatcaaccc taacccaagg ggaacagcct gcctgacagt 60

ggcttgcca cccatgaata cttcctagtc tagtccgttt gtgaaactca gcccatccca 120

acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctct 180

ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg acccctttgg 240

gggaatcaa cgaccctta caggggtcac atatcatcta tcctatatgt caggtattta 300

cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360

gattgaaggt caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg 420

gggaaagggt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca 480

gaaagcctt cagctgttct gctggggctc ttagtaagtc tgagttagaa ctgtatgtac 540

caggtctgct tcttatgggt ggagccaaga cgcatgtgg gtggagcga gacgcaacct 600

caccttctag ctctgcatcc atagcaagta gcctaattgtt tctgtgtcta ggtgtcatct 660

ctgtgaatcg agatccttgg ctttgcttga attagggagg cacaaaatac tcagagattc 720

aagactgctc agcagccag agtccttcct caaaggaaag gtctcaactc tcagcccccc 780
ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg 840
cttgcactt caagaggtca aagaaaatag tgttaaccat gaaaacgaga agaccaacag 900
ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gaggggaacc 960
cacgaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgac tgaagattct 1020
agaaaacacaa tttgctggtt gaacagctga agtgggtgg ggttcttac cccatgttca 1080
tggaagggtg agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca 1140
ccctaattaa cactccctc ttctactgac acccccttca ctctccttct tcataaaaaaa 1200
taaaaaaaagt atttatgtg gctttacga tagaatctt cctcgaacta taaaaagatc 1260
taaatattha tatTTTcac atttaatat ctttagcgatg acaagccaga aacaagtatt 1320
ttttgcctct ctcaacagca aagcttgggg ccttttgggt tccgtgttag gaatagaaca 1380
cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag 1440
acgcacattt ttctcggct ctcttaagct tttcccacag cattggaaa ctttactgac 1500
agcatccaag ttgtgtttct gctaagaact ggactcacat ctctctgtgc atcacttcgg 1560
cccgaaaaatgg ggttagatcct ctgattagcc ttcagatttta gaacacggtg agcctgttgt 1620
gcactaatta tggccagtga caccatagag tcaaagtgcata ttactgaatg ctttcaattt 1680
ctcctaattgc tggtagatgc gcatgtcaca gggccatttt agctgcagac atcactccag 1740
agaattccaa acagatagag acaagtggca cccagaccca tctccttccc ctcggctga 1800
ttatccccag aaataggatg tcccaaagca acactccca gccaactgga gtgctgataa 1860
gtccagttat cagaaagata tggctgtaaat tgtgtatgcac agtgcttgca ttttcttgat 1920
acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980
acaggcagct gtcccttgcc ttcccgatac gtggatgac tcgcattgct gagcggtgt 2040
gtcactgcca aaggaatgac cctctcacat ttcttcctga ttgcatacg ccgcggccag 2100
cttgcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160
ctgaattggc cactgggtgg ggcagggtg tgacttggct tcccaggctg gaagattatc 2220
tcacccagcc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag 2280
gggttcatcc acaagagaga aaaacataga ctcgaggctt agggagctt catgcctgca 2340

ggtcggaggc caccatgg

2358

<210> 2
<211> 2074
<212> DNA
<213> Homo sapiens

<400> 2
ctgcagcaag ttacttaatg tttttgcct cagcatcctc tctgtaaaat gagagcatta 60
gtcttgctcc aacttcgagg gcatggacag ctctggatt tcatatccaa gacccttaaa 120
catcccacag tccttccccc aaacacttct cctcctaata cctccctcag tttgggtcag 180
gcctggaaca aaaaggcata cgaaatggta gaaaaagtgt ccatgactac ttctgactta 240
gatgaagaga ccaatgaaaa tagtaatgac tctgttgct tcagcaggac atatactaaa 300
ataggagcta tacaaagaag attagcatgg actctgtgca agaatgacac acaaatttgt 360
gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaaaag 420
aaaatagtga tagctgtgtc catctcaaag aaaagccag gagatttcct ttatattaccc 480
ccttaagat agaatattag gagaccggaa catatgatac aggaggtact gggagggtcc 540
ctcttgta atgtttgtc ttggggtggg gagtcgatgt cttctcaaag tttcagaaac 600
accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt 660
gggtgagttt ggggagaaaat agacacacaa aggtcaaaca taacttccta attaacactt 720
ccctccattc acaattccct tctccattc ttctctcctg tctttacts akaraaaccc 780
agttttcct gaaactataa aaataccccc agtatgtta cataatttac acctcaaaga 840
tttagaaacca gaaatagaga cttttcaac cttccggaa gcaaagtgca ttatccctcc 900
agccacgtgt ctcaaattttt gatgcatcag aatcatctgg gtgcttkaa attcaagatg 960
attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta 1020
tttctgacaa gctcccacag gtgattcctt tccccacagc atttgagaac ttcagctcaa 1080
tgacctaatc agagtccctgc cattgctaattt atctggtctc attttbtca tatatatata 1140
tagtatttgtt ggttagagatg ggatttgcc atgttgccca ggctagtatt gaactcctaa 1200
gctaagcaat cttccctgtct ctgcctccca aaatgttggg attacagggtg taagccactg 1260



caccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa 1320
gtaaaaatgc tccaattatt atgctgtttt agaacacggt aagcatgtca tgtgctaatg 1380
gccagtgaca tcataaaaaga aaagtgcatt actgaatgct ttcaatgtct tataatgatg 1440
gtaaggtggc atgtcatggg gcctatttag cccagacatc actccaaaga attccaaaca 1500
gatatagaca agtgcctta gggcccagat ccctccct caggctgtt acccagggaa 1560
taggatgtcc tgggacaagt ttccccataag tgaagtgtt gataagtctgc ttatcagaaa 1620
gatattactg ggggtgtat atgttagggca tctacattt cttgataggt agtcatatga 1680
aagctgacaa agaaaaaaag ggcagtgtat tggtgcaatg tcaacagaca gctgtccc 1740
gactcttgc acataggatg acttgcattt ctgagcgatg tgatcaccac caaaggaatg 1800
gcctctcac atttcttcct gattcacata ttcagcaggg ttagcttgc ctccccccc 1860
tcttcagctt cccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920
taatgggggc gggagtgtta cttcggttcc caggttgaa gattatctca cccggcccc 1980
gctatataag ctgaccggtg tggaggggccc cagcagggcc aactccaggg attccttcca 2040
cgacagaaaa acataacaaga ctccctcagc caac 2074

<210> 3
<211> 750
<212> PRT
<213> Homo sapiens

<400> 3
Met Gly Glu Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser
1 5 10 15
Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val
20 25 30
Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val
35 40 45
Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro
50 55 60
Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp
65 70 75. 80
Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp
85 90 95

Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro Ala Ser
100 105 110

Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro His Glu
115 120 125

Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys Gly Asp
130 135 140

Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly Cys Lys
145 150 155 160

Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp Arg Cys
165 170 175

Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys Gln Tyr
180 185 190

Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn Ala Ile
195 200 205

Arg Phe Gly Arg Met Pro Gln Ala Glu Lys Glu Lys Leu Leu Ala Glu
210 215 220

Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp Leu Arg
225 230 235 240

Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser Phe Pro Leu
245 250 255

Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr Thr Asp Lys
260 265 270

Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met Gly Glu Asp
275 280 285

Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln Ser Lys Glu
290 295 300

Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser Val Glu Ala
305 310 315 320

Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly Phe Val Asn
325 330 335

Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly Val His Glu
340 345 350

Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp Gly Val Leu
355 360 365

Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu Lys Ser Leu

370	375	380	
Arg	Lys Pro Phe Gly Asp	Phe Met Glu Pro Lys	
385	390	395	
Lys Phe Asn Ala Leu	Glu Leu Asp Asp	Ser Asp Leu Ala Ile	
405	410	415	
Ala Val Ile Ile Leu Ser	Gly Asp Arg	Pro Gly Leu Leu Asn Val	
420	425	430	
Pro Ile Glu Asp Ile Gln Asp Asn	Leu Leu Gln Ala Leu	Glu Leu Gln	
435	440	445	
Leu Lys Leu Asn His Pro	Glu Ser Ser Gln Leu	Phe Ala Lys Leu Leu	
450	455	460	
Gln Lys Met Thr Asp	Leu Arg Gln Ile Val	Thr Glu His Val Gln Leu	
465	470	475	480
Leu Gln Val Ile Lys	Lys Thr Glu Thr Asp	Met Ser Leu His Pro Leu	
485	490	495	
Leu Gln Glu Ile Tyr Lys Asp	Leu Tyr Ala Trp Ala Ile	Leu Thr Gly	
500	505	510	
Lys Thr Thr Asp Lys Ser	Pro Phe Val Ile Tyr Asp	Met Asn Ser Leu	
515	520	525	
Met Met Gly Glu Asp Lys	Ile Lys Phe Lys His	Ile Thr Pro Leu Gln	
530	535	540	
Glu Gln Ser Lys Glu Val	Ala Ile Arg Ile Phe	Gln Gly Cys Gln Phe	
545	550	555	560
Arg Ser Val Glu Ala Val	Gln Glu Ile Thr	Glu Tyr Ala Lys Ser Ile	
565	570	575	
Pro Gly Phe Val Asn Leu	Asp Leu Asn Asp	Gln Val Thr Leu Leu Lys	
580	585	590	
Tyr Gly Val His Glu Ile	Ile Tyr Thr Met	Leu Ala Ser Leu Met Asn	
595	600	605	
Lys Asp Gly Val Leu Ile	Ser Glu Gly Gln Gly	Phe Met Thr Arg Glu	
610	615	620	
Phe Leu Lys Ser Leu Arg	Lys Pro Phe Gly Asp	Phe Met Glu Pro Lys	
625	630	635	640
Phe Glu Phe Ala Val	Lys Phe Asn Ala Leu	Glu Leu Asp Asp Ser Asp	
645	650	655	

Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly
660 665 670
Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln
675 680 685
Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu
690 695 700
Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr
705 710 715 720
Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met
725 730 735
Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
740 745 750

P<210> 4
D<211> 30
D<212> DNA
D<213> Artificial sequence

P<400> 4
DGGCGATTAA ATAATGTAGT CTTATGCAAT

30

P<210> 5
D<211> 31
D<212> DNA
D<213> Artificial sequence

P<400> 5
DGGGTCTAGA AGGTGCACAC CAATGTGGTG A

31